

Open Research Online

The Open University's repository of research publications and other research outputs

On using a digital resources for the study of an ancient text: the case of Herodotus' *Histories*

Book Section

How to cite:

Barker, Elton; Isaksen, Leif; Rabinowitz, Nick; Bouzarovski, Stefan and Pelling, Chris (2013). On using a digital resources for the study of an ancient text: the case of Herodotus' *Histories*. In: Dunn, Stuart and Mahony, Simon eds. *The Digital Classicist. Bulletin of the Institute of Classical Studies Supplement* (122). Institute of Classical Studies, University of London, pp. 45–62.

For guidance on citations see [FAQs](#).

© 2013 The Institute of Classical Studies

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

THE DIGITAL CLASSICIST 2013

BULLETIN OF THE INSTITUTE OF CLASSICAL STUDIES SUPPLEMENT 122

DIRECTOR & GENERAL EDITOR: JOHN NORTH

DIRECTOR OF PUBLICATIONS: RICHARD SIMPSON

**THE
DIGITAL CLASSICIST
2013**

**EDITED BY
STUART DUNN
AND SIMON MAHONY**

**INSTITUTE OF CLASSICAL STUDIES
SCHOOL OF ADVANCED STUDY
UNIVERSITY OF LONDON**

2013

The cover image is of a torso of Pothos (Roman 1st century BC – 1st century AD) in the Museu Calouste Gulbenkian, Lisbon, Portugal.
Photo © Simon Mahony 2013. All rights reserved.

ISBN 978-1-905670-49-9

© 2013 Institute of Classical Studies, University of London

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the publisher.

The right of the contributors to be identified as the authors of the work published here has been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

Designed and computer typeset at the Institute of Classical Studies.

Printed by Short Run Press Limited, Bittern Road, Exeter EX2 7LW.

This volume is dedicated to the memory of two people whose untimely death marks a great loss, both personally and to our communities.

Elaine Matthews (died 26 June 2011): one of our esteemed contributors, ambassador and advocate of the Digital Humanities and the place there for Classics, thank you for all your many contributions to scholarship, to this volume, and your generous words on the cover of the earlier *Digital Classicist* (Ashgate 2010) volume.

Gerhard Brey (1954-2012): a valued friend, colleague, and collaborator with whom we shared intellectual ideas as well as coffee and biscuits. Gerhard was always willing to seek out new areas of 'interest' and so could be willingly called upon to review chapters in this and the earlier Ashgate volume.

TABLE OF CONTENTS

Acknowledgements	ix
Abstracts	xi
Abbreviations	xv
Introduction	1
Modelling	
Andrew Bevan <i>Travel and interaction in the Greek and Roman world. A review of some computational modelling approaches</i>	3
Vince Gaffney, Phil Murgatroyd, Bart Craenen, and Georgios Theodoropoulos <i>'Only individuals': moving the Byzantine army to Manzikert</i>	25
Texts	
Elton Barker, Leif Isaksen, Nick Rabinowitz, Stefan Bouzarovski, and Chris Pelling <i>On using digital resources for the study of an ancient text: the case of Herodotus' Histories</i>	45
Marco Büchler, Annette Geßner, Monica Berti, and Thomas Eckart <i>Measuring the influence of a work by text re-use</i>	63
Tobias Blanke, Mark Hedges, and Shrija Rajbhandari <i>Towards a virtual data centre for Classics</i>	81
Ryan Baumann <i>The Son of Suda On-line</i>	91
Infrastructure	
Elaine Matthews and Sebastian Rahtz <i>The Lexicon of Greek Personal Names and classical web services</i>	107
Simon Mahony: <i>HumSlides on Flickr: using an online community platform to host and enhance an image collection</i>	125
Valentina Asciutti and Stuart Dunn <i>Connecting the Classics: a case study of Collective Intelligence in Classical Studies</i>	147
Index	161

ACKNOWLEDGEMENTS

The editors would like to thank the Institute for Classical Studies, and specially the Deputy Director and Administrator Olga Krzyszkowska, for their continued support and generosity in hosting and supporting the Digital Classicist seminars. Thanks are also due to all the members of our community who have presented papers at our seminars and conference panels as well as all those who have come along to listen and support these events.

We are grateful to the following scholars for comments, criticism, and advice on individual chapters in this volume; it is through their input that we are able to ensure the high quality of the final work: Chris Blackwell; Gabriel Bodard; John Bodel; Kalina Bontcheva; Gerhard Brey; Hugh Cayless; Graeme Earl; Michael Fulford; Sebastian Heath; Tim Hill; Kathryn Piquette; Dot Porter; Julian Richards; Matteo Romanello; Charlotte Roueché; Melissa Terras; Notis Toufexis; Charlotte Tupman; Michelle Wienhold.

ABSTRACTS

- Andrew Bevan *Travel and interaction in the Greek and Roman World. A review of some computational modelling approaches* pp. 3-24

Inferring dynamic past behaviours from the static archaeological record is always a challenge, but computational and quantitative techniques can be helpful. In particular, they can provide useful insight on patterns of movement and interaction, by better characterising existing archaeological evidence, suggesting simple models of mobile decision-making or proposing expected patterns against which the observed record can be compared. This paper reviews the range of modelling options now available for understanding the movement and interaction behind the archaeological and historical record. There are increasing opportunities not only to pick and choose between different modelling approaches, but also to integrate them in a more theoretically and practically satisfactory way.

- Vince Gaffney, Phil Murgatroyd, Bart Craenen, and Georgios Theodoropoulos
'Only individuals': moving the Byzantine army to Manzikert pp. 25-43

Traditionally, history has frequently emphasized the role of the 'Great Man or Woman', who may achieve greatness, or notoriety, through the consequences of their decisions. More problematic is the historical treatment of the mass of the population. Agent-based modelling is a computer simulation technique that can not only help identify key interactions that contribute to large scale patterns but also add detail to our understanding of the effects of all contributors to a system, not just those at the top. The Medieval Warfare on the Grid project has been using agent-based models to examine the march of the Byzantine army across Anatolia to Manzikert in AD 1071. This article describes the movement model used to simulate the army and the historical sources on which it was based. It also explains why novel route planning algorithms were required in order to surmount problems with standard solutions.

- Elton Barker, Leif Isaksen, Nick Rabinowitz, Stefan Bouzarovski, and Chris Pelling
On using digital resources for the study of an ancient text: the case of Herodotus's 'Histories' pp. 45-62

Involving the collaboration of researchers from Classics, Geography, and Archaeological Computing, and supported by funding from the AHRC, *Hestia* aims to enrich contemporary discussions of space by developing an innovative methodology for the study of an ancient narrative, Herodotus's *Histories*. Using the latest digital technology in combination with close textual study, we investigate the geographical concepts through which Herodotus describes the conflict between Greeks and Persians. Our findings nuance the customary

topographical vision of an east versus west polarity by drawing attention to the topological network culture that criss-crosses the two, and develop the means of bringing that world to a mass audience via the internet.

-In this chapter we discuss three main digital aspects to the project: the data capture of place-names in Herodotus; their visualization and dissemination using the web-mapping technologies of GIS, Google Earth, and Timemap; and the interrogation of the relationships that Herodotus draws between different geographical concepts using the digital resources at our disposal. Our concern will be to set out in some detail the digital basis to our methodology and the technologies that we have been exploiting, as well as the problems that we have encountered, in the hope of contributing not only to a more complex picture of space in Herodotus but also to a basis for future digital projects across the Humanities that spatially visualize large text-based corpora. With this in mind we end with a brief discussion of some of the ways in which this study is being developed, with assistance from research grants from the Google Digital Humanities Awards Program and JISC.

Marco Büchler, Annette Geßner, Monica Berti, and Thomas Eckart

Measuring the Influence of a Work by Text Re-Use

pp. 63-79

Over the centuries an incredible amount of ancient Greek texts have been written. Some of these texts still exist today whereas other works are lost or are available only as fragments. Without considering intentional destruction, one major question remains: why did some texts remain and others get lost? The aim of this chapter is to investigate this topic by trying to determine the influence of certain ancient Greek works through detecting text re-use of these works. Text re-use measures if and how an author quotes other authors and in this chapter we differentiate between *re-use coverage* and *re-use temperature*.

Tobias Blanke, Mark Hedges, and Shrija Rajbhandari

Towards a virtual data centre for Classics

pp. 81-90

A wide variety of digital resources have been created by researchers in the Classics. These tend to focus on specific topics that reflect the interests of their creators; nevertheless they are of utility for a much broader range of research, and would be more so if they could be linked up in a way that allowed them to be explored as a single data landscape. However, while the resources may be reusable, the variety of data representations and formats used militates against such an integrated view. We describe two case studies that address this issue of interoperability by creating virtual resources that are independent of the underlying data structures and storage systems, thus allowing heterogeneous resources to be treated in a common fashion while respecting the integrity of the existing data representations.

Ryan Baumann *The 'Son of Suda On-line'*

pp. 91-106

The Son of Suda On-Line (SoSOL) represents the first steps towards a collaborative, editorially-controlled, online editor for the Duke Databank of Documentary Papyri (DDbDP). Funded by the Andrew W. Mellon Foundation's Integrating Digital Papyrology Phase 2

(IDP2), SoSOL provides a strongly version-controlled front-end for editing and reviewing papyrological texts marked up in EpiDoc XML.

Elaine Matthews and Sebastian Rahtz

The Lexicon of Greek Personal Names and classical web services pp. 107-24

This chapter documents the data resources of the long-term classical research project, *The Lexicon of Greek Personal Names* (LGPN), published in six volumes since 1987. It explains and demonstrates the web interfaces and services which now make available online the bulk of the LGPN, providing both powerful searching tools for scholars and an interface to allow other systems to link to LGPN data. Making the data available online provides direct, unmediated access to the material and supports exploitation of the data for further research both individual and collaborative.

We describe the work that went into creating the Lexicon, detail the granularity of the data structures, and explain the history of the project's record management. We then move onto the work undertaken in recent years to provide an archival XML-based format for the Lexicon's long-term preservation, and show how this has allowed us to build new web services, including exposure of Resource Description Framework (RDF) metadata, using the ontology of the CIDOC Conceptual Reference Model (CRM) ontology for semantic web applications.¹

Simon Mahony: *HumSlides on Flickr: using an online community platform to host and enhance an image collection*

125-46

Moving a teaching and research image collection from an analogue to a digital medium for delivery brings with it many advantages but at the same time it also presents many new problems and ones probably not previously considered. This chapter discusses the move of a departmental slide collection, firstly to a proprietary in-house format, and then subsequently to the online community platform Flickr. It draws on the experience and model of the Library of Congress in partnership with Flickr and *The Commons*, as well as initiatives at Oxford and at New York University, and in doing so critically analyses and evaluates the possibilities for the future development of this collection. It asks why this collection is not currently being used to its potential and examines how the development of a user community would help to enrich the collection and ensure long term sustainability and future growth.

¹ CIDOC CRM is an ISO standard (21127:2006) that 'provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation' <<http://www.cidoc-crm.org/>>.

Valentina Asciutti and Stuart Dunn

Connecting the Classics: a case study of Collective Intelligence

in Classical Studies

pp. 147-60

One of the great potentials of the internet is its capacity to aggregate and unify information from diverse sources. Information in the Classics, and data generated by classicists, is inherently fragmented, and organized according to different standards. This paper describes a project at King's College London which sought to provide a set of aggregating services to humanities scholars. www.arts-humanities.net provides a platform, a library, and a taxonomy to organize and present data: we describe its facilities for supporting a multi-source dataset tracing the paths of Romano-British inscriptions, both in space and conceptually. Itinerant geographies of metrical versus text inscriptions are discussed, including how these can be published in a variety of non-conventional platforms, such as Twitter. We argue that, in the future, these platforms will come to play a critical role in the wider scholarly discourse of the Classics.

ABBREVIATIONS

ABM	Agent-based modelling
ADS	Archaeology Data Service
AHRC	Arts and Humanities Research Council
API	Application Programming Interface
APIS	Advanced Papyrological Information System
AWIB	Ancient World Image Bank
CC	Creative Commons
CI	Collective Intelligence
CIDOC	International Council of Museums
CRM-CIDOC	CIDOC Conceptual Reference Model
CSV	Comma-separated data fields
DANS	Data Archiving and Networked Services
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DDbDP	Duke Databank of Documentary Papyri
DPI	Dots per inch
DVCS	Distributed Version Control Systems
EDM	Europeana Data Model
GAP	Google Ancient Places
GIS	Geographical Information Systems
HEA	Higher Education Academy
HEFCE	Higher Education Funding Council for England
HESTIA	Herodotus Encoded Space-Text-Image Archive
HGV	<i>Heidelberg Gesamtverzeichnis der griechischen Papyrusurkunden Ägyptens</i>
Iaph	<i>Inscriptions of Aphrodisias</i>
IDP	Integrating Digital Papyrology
ISAW	Institute for the Study of the Ancient World
JDI	Image Digitization Initiative
JISC	Joint Information Systems Committee
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
LaQuAT	Linking and Querying Ancient Texts
LCCW	Longest Common Consecutive Words
LGPN	<i>The Lexicon of Greek Personal Names</i>
LoC	The Library of Congress
MDID	Madison Digital Image Database
OAI-ORE	Open Archives Initiative Object Reuse and Exchange
OER	Open Education Resources
OGSA-DAI	Open Grid Service Architecture–Data Access and Integration
OGSA-DQP	Distributed Query Processing
PDF	Portable Document Format

PN	Papyrological Navigator
PRM	Probabilistic Road Map
RDF	Resource Description Framework
RIB	<i>Roman Inscriptions of Britain</i>
SGML	Standard Generalized Markup Language
SOL	Suda On-Line
SoSOL	Son of Suda On-Line
SQL	Structured Query Language
SVG	Scalable Vector Graphic
TEI	Text Encoding Initiative
TLG	<i>Thesaurus Linguae Graecae</i>
URI	Uniform Resource Identifiers
V&A	Victoria and Albert Museum
VRE	Virtual Research Environment
WFS	Web Feature Service
WMS	Web Map Service
WYSIWYG	What-You-See-Is-What-You-Get
XML	Extensible Markup Language

ON USING DIGITAL RESOURCES FOR THE STUDY OF AN ANCIENT TEXT: THE CASE OF HERODOTUS' *HISTORIES**

ELTON BARKER, LEIF ISAKSEN, NICK RABINOWITZ,
STEFAN BOUZAROVSKI, and CHRIS PELLING

1. *Approaching the world of Herodotus*

About a generation after the unexpected and remarkable defeat of the Persians by a ragtag assembly of Greek city-states, Herodotus of Halicarnassus set out to investigate the great deeds of these two peoples and why they had come into conflict in the first place. His enquiry took him all round the Mediterranean, 'traversing in detail towns of men both small and great alike, for', as he explains, 'of the places that were once great, most have now become small, while those that were great in my time were small before' (ὁμοίως σμικρὰ καὶ μεγάλα ἄστεα ἀνθρώπων ἐπεξιών· τὰ γὰρ τὸ πάλαι μεγάλα ἦν, τὰ πολλὰ σμικρὰ αὐτῶν γέγονε· τὰ δὲ ἐπ' ἐμεῦ ἦν μεγάλα, πρότερον ἦν σμικρὰ: 1.5). Previous scholarship has typically depicted that world as being divided into three separate units, Europe, Asia, and Libya, and has highlighted the importance of water bodies, in particular rivers, for organizing that space.¹ Such 'mental maps' (see Figure 1)² are useful for drawing attention to the social, political, and cultural constructions of space,³ but are

* We are extremely grateful to the Arts and Humanities Research Council for funding this research. The initial funding (for Barker's Early Career grant) ran from September 2008 to June 2010. In July 2013 the AHRC awarded Hestia 'Follow on' funding for one year in order to consolidate the technologies and promote knowledge exchange with groups within and outside academia. An earlier account of this work, entitled 'Mapping an ancient historian in a digital age: the Herodotus Encoded Space-Text-Image Archive (HESTIA)', is published by the *Leeds International Classical Journal* 9.1 (2010): (<<http://www.leeds.ac.uk/classics/lics/>>). All translations of Herodotus are our own.

¹ For the importance of natural boundaries to Herodotus' conception of history and, in particular, rivers as demarcation boundaries and markers of transgression, see esp. H. Immerwahr, *Form and thought in Herodotus* (Cleveland 1966), in the index under *river motif*; cf. D. Braund, 'River frontiers in the environmental psychology of the Roman world', in *The Roman army in the east*, ed. D. L. Kennedy, *JRA* Supp 17 (Ann Arbor 1996) 43-47.

² This map has been chosen because, published under a Wikimedia Commons licence, it can be reproduced without any infringement of copyright. But it also typifies other reproductions of Herodotus' world: see O. A. W. Dilke, *Greek and Roman maps* (London 1985) 58.

³ See e.g. N. J. W. Thrower, *Maps and civilization: cartography in culture and Society* (Chicago 1996); J. B. Harley, 'Deconstructing the map', *Cartographica: The International Journal for Geographic Information and Geovisualization* 26 (1989) 1-20; P. Jackson, *Maps of meaning: an*

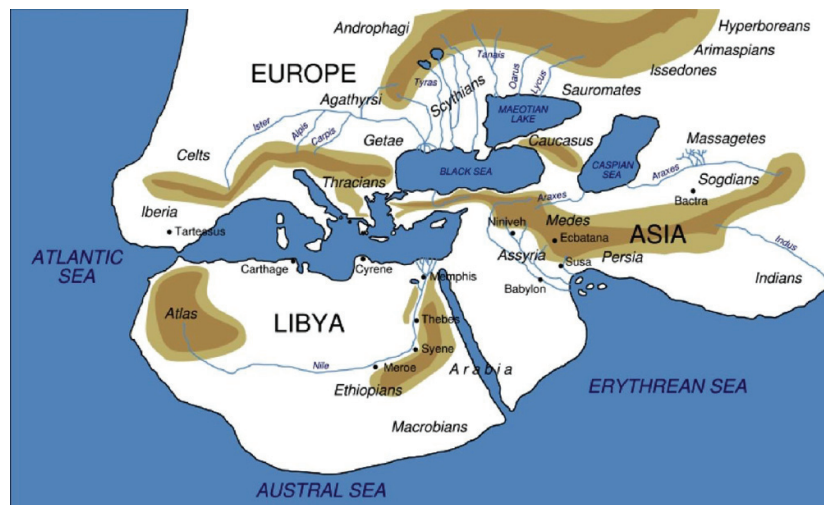


Figure 1: Herodotus' world (source: http://en.wikipedia.org/wiki/File:Herodotus_world_map-en.svg, Last accessed 15.11.2013)

inevitably hampered by the medium of representation. The printed map can offer only a compressed snapshot of how space is conceived across Herodotus' narrative as a whole, and tends to define it, moreover, in terms of separate territories, whose boundaries are rigidly policed by rivers, seas, and the odd desert. Herodotus himself appears to be deeply sceptical of the capacity for maps – a new technology at the time of the author – to represent adequately the kinds of engagement between peoples that he depicts in his narrative.⁴ Our own technological advances, however, are now creating the possibilities not only of relating space more systematically and sensitively to Herodotus' text, but also of broadcasting the results to a much wider community. In short, the digital medium offers the potential to capture a world in flux, which changes over time (as Herodotus recognizes in the quotation above), and to represent places within that world in relation to each other.

Involving the collaboration of researchers from Classics, Geography, and Archaeological Computing, Hestia has been experimenting with various ways of extracting and analysing the different ways in which geographic space is organised in Herodotus' narrative.⁵ Using the latest digital technologies in combination with close textual study, we examine the

introduction to cultural geography (London 1994); K. T. Jones, 'Scale as epistemology', *Political Geography* 17 (1998) 25-28.

⁴ Hdt 4.37. Rather than talking in terms of abstract conceptions of space, Herodotus narrates the geography from the perspective of one travelling: *i.e.* the order of description does not necessarily map onto the nearest place but the one that comes next, as one travels to it. For 'hodological' approaches to space in Herodotus, see A. Purves, *Space and time in ancient Greek narrative* (Cambridge 2010) 144-49; *cf.* P. Janni, *La Mappa e il Periplo. Cartografia antica e spazio odologico* (Marcerata 1984); J. S. Romm, *The edges of the earth in ancient thought: geography, exploration, and fiction* (Princeton 1994) 34-44.

⁵ For more information, please go to the project website: <<http://hestia.open.ac.uk/>>.

geographical concepts through which Herodotus describes the conflict between Greeks and Persians,⁶ and explore the connections between places that Herodotus makes over the course of his narrative. Our research aims are, first, to question the received Greek-other dichotomy of the *Histories* by revealing ambiguities in the ways in which Greek and Persian worlds are mapped out and interrelated,⁷ and, second, to use a network analysis for thinking about Herodotus' textualization of space – how, in other words, the narrator puts spatial ideas and concepts into words.⁸ In both cases, we experiment with various kinds of network graph to investigate the topological relationships between places rather than try to reconstruct a topographical reality.

For the purposes of this volume we discuss three main digital aspects to the project: the data capture of place-names in Herodotus; their visualization and dissemination using digital technologies like Geographical Information Systems (GIS)⁹ and web-mapping applications such as Google Earth and Timemap.js;¹⁰ and the interrogation of the relationships that Herodotus draws between different geographical concepts using the digital resources at our disposal. Our concern here will be to set out in some detail the digital basis to our methodology and the technologies that we have been exploiting, as well as the problems that we have encountered, in the hope of contributing not only to offering a more complex picture of space in Herodotus but especially to setting out a basis for future digital projects across the humanities that deal with the spatial representation of large text-based corpora.¹¹ With

⁶ In a recent article T. Harrison, 'The place of geography in Herodotus' *Histories*', in *Travel, geography and culture in ancient Greece, Egypt and the Near East*, ed. C. Adams and J. Roy (Oxford 2007) 44-65, makes the 'search of geography and its place in the *Histories*' (44) the highest priority.

⁷ E.g. E. Hall, *Inventing the barbarian* (Oxford 1989); P. A. Cartledge, *The Greeks and others* (Bristol 2002 [1993]). The 'East vs. West' dichotomy in Herodotus has recently been challenged, however: see for example, Chris Pelling, 'East is east and west is west – or are they?', *Histos* 1997 51-66: <<http://research.ncl.ac.uk/histos/documents/1997.04PellingEastIsEast5166.pdf>> (revised and updated with further bibliography, in *Herodotus volume 2: Oxford readings in Herodotus*, ed. R. V. Munson (Oxford 2013) 360-79).

⁸ Network theory in the field of ancient Greek history: I. Malkin, *The returns of Odysseus* (Berkeley 1998); C. Constantakopoulou, *The dance of the islands* (Oxford 2005). The two-dimensional 'Cartesian'-style map has dominated Western horizons since the Enlightenment: A. J. Gurevich, *Categories of Medieval culture* (London 1985). For a fuller articulation of the geographic theory underpinning our analysis of Herodotean space, see E. T. E. Barker, S. Bouzarovski, C. B. R. Pelling, and L. Isaksen, ed., *New worlds out of old texts: developing techniques for the spatial analysis of ancient narratives* (Oxford forthcoming).

⁹ According to Wikipedia (<http://en.wikipedia.org/wiki/Geographic_information_system>), Geographic information system (GIS) is 'a system designed to capture, store, manipulate, analyze, manage and present all types of geographical data [last accessed on 15/11/2013]. In short, GIS allows the researcher to represent spatial data visually – to think with maps, as it were.

¹⁰ 'Timemap.js is a Javascript library to help use online maps, including Google, OpenLayers, and Bing, with a SIMILE timeline': <https://code.google.com/p/timemap/> [last accessed on 15/11/2013].

¹¹ Developments in technology are fast revolutionizing research into and visualization of geographical concepts, all of which has the potential to transform customary notions of mapping space in humanities more broadly: see, for example, R. J. A. Talbert, 'Greek and Roman mapping: twenty-first century

```

132 <milestone n="1" unit="chapter"/><milestone n="0" unit="section"/>
133 <milestone unit="para"/>This is the display of the inquiry of <name
134 type="pers">Herodotus</name> of <placeName
135 key="tgn,7016142">Halicarnassus</placeName>, so that things done by man not be
136 forgotten in time, and that great and marvelous deeds, some displayed by the <name
137 type="ethnic">Hellenes</name>, some by the barbarians, not lose their glory,
138 including among others what was the cause of their waging war on each other.
139 <milestone n="1" unit="section"/>
140 <milestone unit="para"/>The <name type="ethnic">Persian</name> learned men say that
141 the <name type="ethnic">Phoenicians</name> were the cause of the dispute. These (they
142 say) came to our seas from the sea which is called Red,<note anchored="true"
143 res="red">Not the modern <placeName key="tgn,7016791">Red Sea</placeName>, but the
144 <placeName key="tgn,7016761">Persian Gulf</placeName> and adjacent waters.</note>
145 and having settled in the country which they still occupy, at once began to make long
146 voyages. Among other places to which they carried <name type="ethnic">Egyptian</name>
147 and <name type="ethnic">Assyrian</name> merchandise, they came to <placeName
148 key="perseus,Argos">Argos</placeName>,

```

Figure 2: TEI P4 XML encoding of Herodotus – see note 12 (source: Perseus)

this in mind, we conclude with a brief discussion of some ways in which this study is being developed, with assistance from Research Grants from Google and JISC, which is taking this research (we hope) into exponentially larger data sets, such as Google Books.

2. Capturing the Data

Given our blended methodology of using digital tools to complement textual analysis, we first needed to acquire or generate a digital version of Herodotus' *Histories*. Fortunately help was at hand in the form of the *Perseus Digital Library*, which not only makes texts freely accessible, but also, by virtue of releasing them under a Creative Commons Attribution Non-Commercial Share-Alike licence, allows them to be used and adapted in whatever way users, such as ourselves, want.¹² We further benefited from the fact that the 'placenames' or *toponyms* – the data which we were interested in extracting from Herodotus – had already been marked up in the English edition of the text according to Text Encoding Initiative (TEI) compliant XML in a semi-automated process (see Figure 2).

While the acquisition of the digital text from *Perseus* gave our project an initial boost, a number of issues resulting from that inheritance were raised, which had to be overcome before the data could be properly stored in a database and used. First, the text had to be updated from the P4 TEI schema used by *Perseus* to the latest version, P5, which we accomplished with minimal information loss using an automated conversion tool

perspectives', in *Cartography in antiquity and the middle ages: fresh perspectives, new methods*, ed. R. J. A. Talbot and R. W. Unger (Leiden 2008) 9-27; and most recently, M. Dear, J. Ketchum, S. Luria and D. Richardson, ed., *GeoHumanities: art, history, text at the edge of place* (New York 2011). Cf. J. B. Harley, 'Deconstructing the map', *Cartographica: The International Journal for Geographic Information and Geovisualization* 26 (1989) 1-20.

¹² For more information about Perseus, see <http://www.perseus.tufts.edu/hopper>, adding /opensource for their licensing policy, /publications for related scholarship, and /dltext?doc=Perseus%3Atext%3A1999.01.0126 to access the text that we used—the 1920 Loeb translation of A. D. Godley.

developed by Sebastian Rahtz.¹³ Second, we settled on using the English text of Herodotus' *Histories* for investigating spatial data, primarily on the basis that the alternative would have meant going through the text and tagging the *toponyms* 'by hand', when they were already widely available to us from *Perseus* using the English version, but also with a view to disseminating the results as widely as possible. Bearing in mind the scholarly importance of the Greek text, however, we assigned each section of the text in both documents (English and Greek) a common identifier in order to draw an association between them.¹⁴ These sections, assigned by the original TEI encoders, equate more or less to a sentence of text, thereby giving a finer level of granularity than the chapter citations traditionally used. At the same time, each section was assigned to its canonical citation, in order to make identification at this level equally possible. Third, we had to strip out all footnotes, which came embedded in the *Perseus* Loeb edition, in order to prevent contamination by modern *toponyms* within them. These 'reference free' documents form the basis for the analysis.

Lastly, many of the 'placenames' inherited from *Perseus* came already tagged with certain kinds of data, such as geographical co-ordinates, modern-day location name, and place type, deriving from both the *Perseus Gazetteer* and the *Getty Thesaurus of Geographic Names*. The information was, however, automatically generated, and thus of a variable nature and sometimes wrong – different identifiers were used for the same location while sometimes places were misidentified altogether. In order to create a consistent basis for further work, inconsistencies and inaccuracies needed to be corrected by hand, specifically: removing duplicate entities for places, correcting references to false places (in particular those with homonyms), and updating co-ordinates. Furthermore, in order to filter and query places and references more effectively so that different place types could be handled separately, we also introduced categories to distinguish between *toponyms* in the database on the basis of their being:

- 1) identifiable communities, encompassing both Greek *poleis* and other kinds of habitations, both Greek and non-Greek (= 'settlements');
- 2) larger areas which may contain a variety of communities within an approximate region (= 'territories');
- 3) natural features of the environment (= 'physical features').

The whole process was labour intensive, but most efficient when the principal academic investigator (Barker) and digital specialist (Isaksen) worked together, allowing for targeted solutions to be developed for clearly defined problems. Additionally, we found the visualization technologies described below, particularly Google Earth, especially useful for editing problem locations, particularly when combined with more traditional reference materials such as the *Barrington Atlas*.¹⁵

¹³ TEI P4 to P5 conversion tool: <<http://www.tei-c.org/Guidelines/P5/p4top5.xsl>>.

¹⁴ That is to say, although we are using the English text to extract spatial data, at every point in the process the Greek text is available for scrutiny alongside it. We also transformed the Greek text from Betacode to Unicode using the Transcoder tool developed by Hugh Cayless, which is available for download via the Open Source directory SourceForge: <<http://sourceforge.net/projects/epidoc/files/transcoder/>>.

¹⁵ R. J. A. Talbot, ed., *Barrington atlas of the Greek and Roman world* (Princeton 2000).

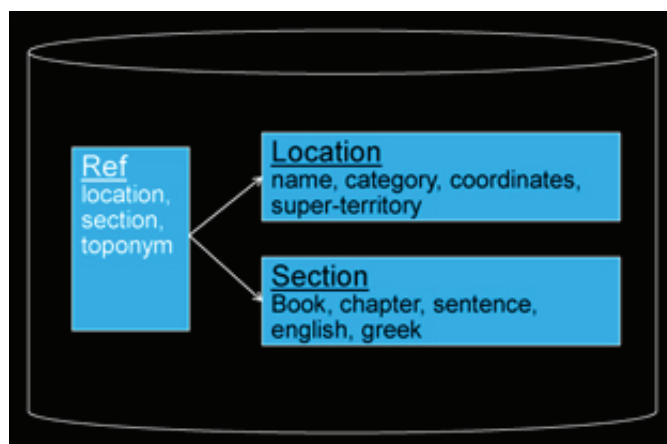


Figure 3: The Hestia database: PostgreSQL

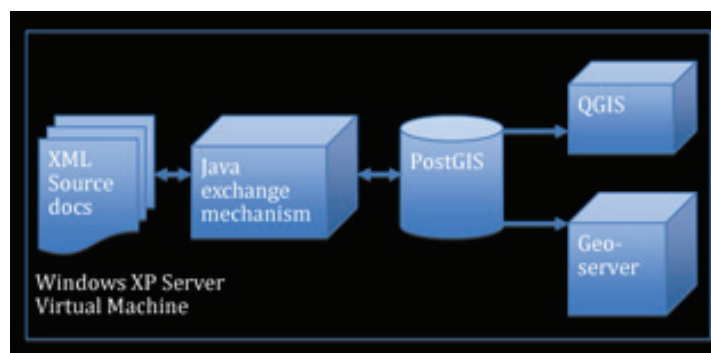


Figure 4: Hestia's architecture

As part of the data cleaning and categorization process, we exported the spatial references to a database, structured simply by (see Figure 3): a *Section* table containing information about the portion of Herodotus' text in which the locations occur, a *Location* table identifying each unique place, and a *Ref* table providing a unique ID for all references to spatial locations within the text.¹⁶ We chose the database type PostgreSQL,¹⁷ an industrial-strength open-source database, on the basis that its PostGIS¹⁸ extension

¹⁶ The advantage of using a structure of this nature is that updates to the Section or Location tables automatically filter through to their References. A virtual table (called a 'view') combining all these together can thereby be created which is accessible to the webmapping server and GIS application.

¹⁷ PostgreSQL is a powerful, open source object-relational database system that has standard compliance: <<http://www.postgresql.org/about/>>.

¹⁸ According to the website <<http://postgis.refrains.net/>>, 'PostGIS adds support for geographic objects to the PostgreSQL object-relational database. In effect, PostGIS 'spatially enables' the

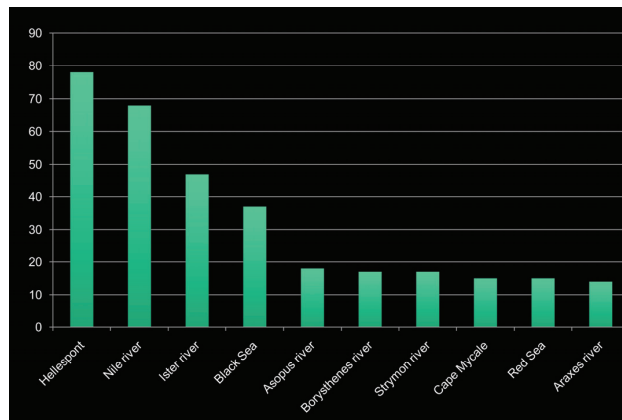


Figure 5: Bar chart showing the ten physical places mentioned most often in the *Histories*

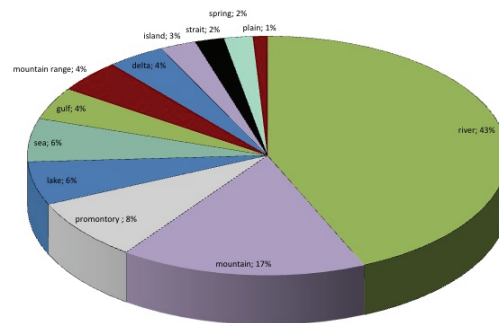


Figure 6: Pie chart showing the proportions according to physical type mentioned in the *Histories*

provides excellent functionality for spatial data and is widely supported by other applications. In particular, we had in mind that, by storing information about references, locations, and the text, it would be possible to link the database to both a Desktop GIS system and Webmapping server simultaneously (see Figure 4).

Just armed with this database, one can now aggregate, query, and analyze Herodotus' spatial data quickly and easily in numerous ways, as demonstrated by Figures 5 and 6. Figure 5 depicts the results of an enquiry into the physical places most frequently mentioned, represented in a tabular form, while Figure 6 shows a pie-chart of the distribution of physical types in the *Histories*. But it is with GIS technology and webmapping applications that we wished to experiment, primarily because of their potential to bring together different kinds of data in a visual representation of narrative. Yet, as we shall see, there is a marked difference in the *usability* of these resources, which, we believe, will prove a significant factor in their

PostgreSQL server, allowing it to be used as a backend spatial database for geographic information systems (GIS)' [last accessed on 15/11/2013].

adoption in the humanities more broadly. As well as outlining the new uses to which Herodotus' spatial data can be put, the rest of this chapter will also highlight limitations in these resources for humanities researchers, not least the considerable investment of time needed for non-IT specialists to become accustomed to the technology. Indeed, it is important to remember that all digital approaches come with their limitations as well as opportunities: we strongly believe that is only through a blended methodology – merging digital techniques with more traditional scholarship – that interesting new insights may emerge.

3. *Thinking about the Histories with GIS*

With this dataset now pre-processed, we were able to apply a variety of technologies to it in order to examine spatial phenomena in Herodotus' *Histories*, primary among which was using GIS. Already well established in archaeology and the social sciences, Geographic Information System (GIS) technology is increasingly being applied to historical studies,¹⁹ though it is widely acknowledged that 'the humanities raise fundamental epistemological and ontological issues for GIS applications', in particular with regard to the contingent and qualitative nature of much of humanities research.²⁰ To the best of our knowledge, applying GIS to the study of an ancient historical narrative marks a distinctive new direction in the use of this technology, and suggests fruitful new lines of enquiry, as well as raising a number of additional challenges, which we set out in sections 4-6 below.²¹

As far as *Hestia* was concerned, GIS enabled us to approach Herodotus' narrative in an innovative and thought-provoking way. While the database may be used to filter records and even create new data, its information can only be delivered internally in table form; GIS, on the other hand, allows us to visualize Herodotus' spatial references more memorably, on a map. Two points about these maps are worth emphasizing from the outset. First, all the maps that we have produced on this project can be reproduced by anyone else: in fact, users are free to generate the maps of *their* choice, a possibility that we have sought to facilitate by using open source GIS applications ourselves.²² Second, all maps generated should be regarded *not* as products demonstrating the world according to

¹⁹ See especially I. N. Gregory and P. S. Ell, *Historical GIS: technologies, methodologies and scholarship* (Cambridge 2007); I. N. Gregory and R. G. Healey, 'Historical GIS: structuring, mapping and analyzing geographies of the past', *Progress in Human Geography* 31 (2007) 638-53; A. K. Knowles, ed., *Placing history: how maps, spatial data, and GIS are changing historical scholarship* (Redlands 2008).

²⁰ See e.g. T. M. Harris, S. Bergeron and L. J. Rouse, 'Humanities GIS: place, spatial storytelling, and immersive visualization in the humanities', in *GeoHumanities*, ed. Dear *et al.* (n. 11 above) 227-40 (quotation on 228).

²¹ Combining GIS with narrative analysis is discussed by Mei-Po Kwan and Guoxiang Ding, 'Geo-narrative: extending Geographic Information Systems for narrative analysis in qualitative and mixed-method research', *The Professional Geographer* 60.4 (2008) 443-65.

²² We use the open source application QGIS (<<http://www.qgis.org/>>), on the basis that anyone else wishing to query our data can use it. It also provides functionality for importing NASA's free Web Mapping Service (WMS) called the Blue Pearl Mosaic, thereby providing a helpful raster backdrop of the Earth's surface, on which Herodotus' spatial data can be projected.

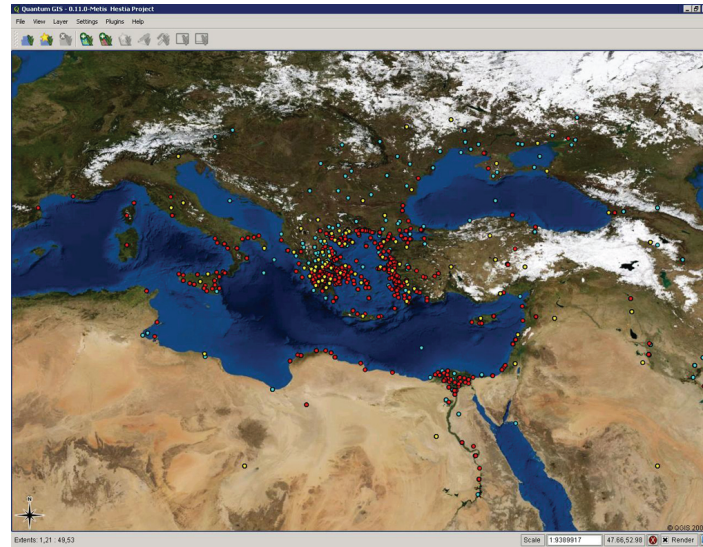


Figure 7: Map (in QGIS) showing all settlements (red), territories (yellow) and physical features (blue) in Herodotus' Histories

Herodotus but as tools of the enquiry: in other words, unlike the customary use of maps as ideological representations of various kinds of spatial data, all *Hestia* maps are interrogative, flagging up in visually arresting ways potential areas of interest that demand analysis and further study.²³

This fundamental point should be clear from the most basic maps that we can generate. Figure 7 simply represents a 'flat' image of the spatial data: it marks all the places that Herodotus mentions over the course of his work with a single dot, thereby providing a snapshot of the huge scope of his enquiry. Yet, such a view gives an idea only of the spatial intensity of the distribution (around the Aegean), not of the frequency of any given place. However, since each place is given a unique identifier in the database that ties it to a particular point in the narrative, it is also possible to generate a count of the references and visualize the results according to a graded symbology. For example, Figure 8 depicts those places categorized as 'settlements', scaled according to the number of references each receives over the course of the *Histories*: the largest circles show those settlements mentioned most often. Of these, Athens and Sparta hardly come as a surprise, being the two most important Greek city-states. Sardis is the place from which Herodotus launches his narrative, and represents an important signifier of growing Persian influence. Arguably Delphi is equally important as a cultural and political centre in the Hellenic world;

²³ As Herodotus himself was acutely aware, the medium in which space is represented can fundamentally alter the way that it is conceived and understood, which is no less true of our own latest satellite imaging technologies in spite (or because) of their apparent claim to accuracy – hence the importance of conceiving these maps as part of the investigative process, not as its end.

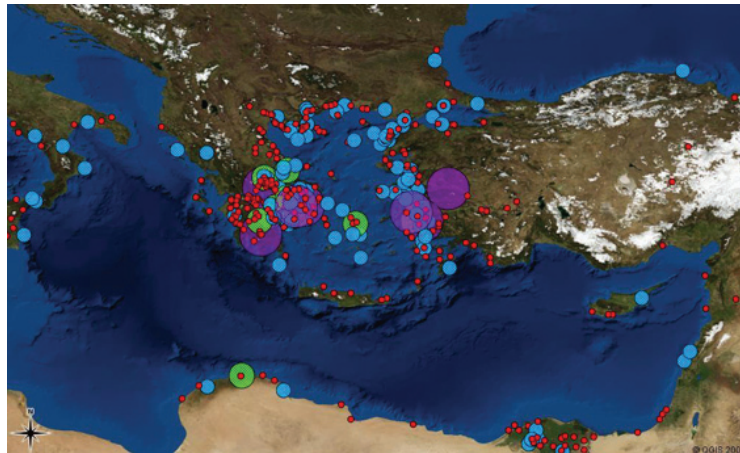


Figure 8: Map (in QGIS) showing settlements according to their number of citations in the *Histories*. The places most frequently mentioned (represented by the large purple circles) are: Sparta, Delphi, Athens, Salamis, Samos, Miletus and Sardis

anyway, there is little doubt of its importance in Herodotus' narrative.²⁴ The same is obviously true of the island of Salamis, the place of the decisive defeat of the Persian fleet by the Greek coalition. The value of this map arguably, then, lies in the attention that it draws to Miletus and Samos: Miletus, as the centre of the Ionian revolt, is, in Herodotus' words, 'the beginning of all evils for Greeks and barbarians alike' (ἀρχὴ κακῶν Ἑλληνσί τε καὶ βαρβάροις: 5.97.3);²⁵ Samos, on the other hand, appears to be important because of the text's post-history and the growth of the Athenian empire, which had resulted in – known to both Herodotus and his audience – the brutal subjugation of Samos.²⁶

²⁴ See: H. W. Parke and D. E. W. Wormell, *The Delphic oracle* (Oxford 1956); R. Crahay, *La littérature oraculaire chez Hérodote* (Paris 1956); H. Flower, 'Herodotus and Delphic traditions about Croesus', in *Georgica: Greek studies in honour of George Cawkwell*, ed. M. A. Flower and M. Toher, *BICS Supp.* 58 (London 1991) 57-77; L. Maurizio, 'Delphic oracles as oral performance: authenticity and historical evidence', *ClassAnt* 16 (1997) 308-34; E. T. E. Barker, 'Paging the oracle: interpretation, identity and performance in Herodotus' *History*', *Greece & Rome* 53 (2006) 1-28; J. Kindt, 'Delphic oracle stories and the beginning of historiography: Herodotus' *Croesus logos*', *Classical Philology* 101 (2006) 34-51. See also Purves, *Space and time* (n. 4 above), 150-58.

²⁵ Herodotus is referring to the ships that the Athenians send to Miletus in support of the revolt. On the significance of this phrase, which comes from Homer (*Il.* 5.62-3), see the contributions of R. V. Munson, 'The trouble with the Ionians: Herodotus and the beginning of the Ionian revolt (5.28-38.1)', 146-67 (esp. 149-59); C. B. R. Pelling, 'Aristagoras (5.49-55, 97)', 179-201 (esp. 182, 186); and J. Henderson, '"The fourth Dorian invasion" and "the Ionian revolt"', 289-310 (esp. 305), all of which are in: *Reading Herodotus. A study of the logoi in Book 5 of Herodotus' Histories*, ed. E. Irwin and E. Greenwood (Cambridge 2007).

²⁶ See E. Irwin, 'Herodotus and Samos: personal or political?', *Classical World* 102 (2009) 395-416; cf. P. Stadter, 'Herodotus and the Athenian *Arche*', *Annali della Scuola Normale Superiore di Pisa* 22 (1992) 781-809; J. Moles, 'Herodotus warns the Athenians', *Papers of the Leeds International*

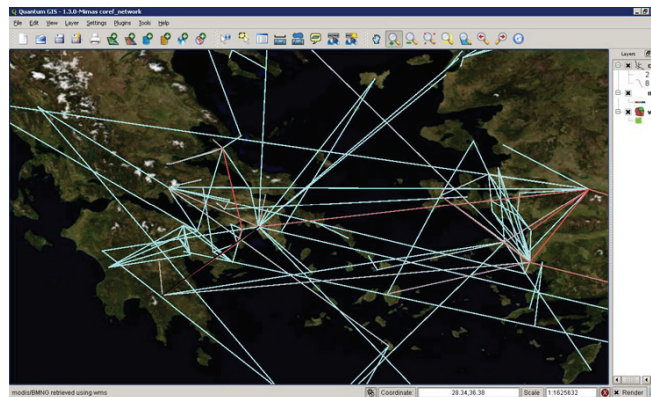


Figure 9: Close-up of map (in QGIS) showing the ‘settlement network’ across the *Hestias*: lines indicate places mentioned together in the text on more than one occasion; those in red indicate the most intense relationships based on count alone

4. Networking places

Both of these kinds of map depict places in isolation from each other: as we outlined above, however, we have been especially keen to approach the question of Herodotus’ narrativization of space through the lens of ‘*topological*’ connections between places – those links that the narrator draws between different spatial concepts as he tells his story. With this in mind, we undertook a close textual and qualitative-based analysis of spatial concepts across one stretch of narrative (Book 5).²⁷ Since, however, this in-depth study was extremely time consuming and labour intensive, we also trialled the extent to which it was possible to use the *Hestia* database and GIS technology as a short-cut to highlighting potential patterns of interest.²⁸

These chiefly included the connections between places that Herodotus draws in his narrative, for which we used a SQL query of the database based on simple co-presence of terms within the same section of text, and visualized the results in QGIS. Since these links were fully automated and we had no idea of telling what the connections were, beyond the fact that two or more places were being mentioned in ‘the same breath’, we tried to minimise arbitrariness by counting only those links that occurred more than once. Still, the resulting picture can be rather chaotic, as seen in Figure 9, which depicts the rough

Latin Seminar 9 (1996) 259-84. On the importance of Samos more generally, including the use of data from *Hestia*, see C. B. R. Pelling’s 2010 Barron Memorial Lecture, ‘Herodotus and Samos’, *BICS* 54 (2011) 1-19.

²⁷ For a discussion of the methodology and results, see: E. T. E. Barker and S. Bouzarovski, ‘Developing a qualitative network analysis: tools, methodological issues and preliminary results’, in *New worlds*, ed. Barker *et al.* (n.8 above).

²⁸ For an example of using GIS in network analysis, see: M. L. Berman, ‘Boundaries or networks in historical GIS: concepts of measuring space and administrative geography in Chinese history’, *Historical Geography* 33 (2005) 118-33.

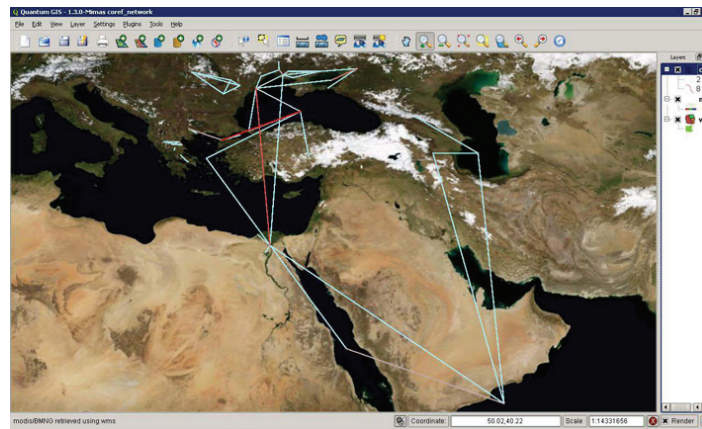


Figure 10: Map (in QGIS) showing the ‘physical features’ network across the *Histories* – those features mentioned together in the text on more than one occasion. A red cross identifies the two strongest links, one between the Ister and Nile, the other between the Hellespont and the Pontus.

‘network’ picture of settlements across the entire *Histories* (focused on the Aegean Sea). It should be remembered, however, that in the GIS application the network can be tested: the user can click on any one these links and bring up the relevant passages in which the relationship is mentioned. Such a map is best thought of as a stimulus for the analysis and closer reading of the text, rather than a representation of (or replacement for reading) it.²⁹

To give a brief example, consider figure 10, which depicts the physical features mentioned together in the text on more than one occasion. Two relationships stand out in red. The first is the connection between the river Ister (the modern-day Danube) in Scythia and the Nile. The strength of this relationship relates, not to a ‘real-life’ network of trade exchange, but rather to a conceptual network utilized by Herodotus to organize the world he narrates. Scythia and Egypt, on the margins of the known world, function as places of comparison through which Herodotus brings to light important aspects of cultural and political identity, in particular in relation to the Greeks and Persians.³⁰ The other strong connection marked in this map is between the Pontus and the Hellespont. This axis is important because it relates to the growing stretch of the Persian Empire, as Darius casts his eyes westwards from the Pontus towards the Hellespont, the bridging point between east and west, in a move that anticipates his son’s literal bridging of the Hellespont in his invasion of Greece.³¹ The map also demonstrates the extent to which

²⁹ SQL or ‘Structured Query Language’ is ‘a special-purpose programming language designed for managing data held in a relational database’: <http://en.wikipedia.org/wiki/SQL> [last accessed on 15/11/2013].

³⁰ F. Hartog, *The mirror of Herodotus* (Berkeley 1988 [1980]). At 2.33-34 Herodotus explicitly draws the comparison between the Ister and Nile.

³¹ On the importance of the Black Sea in Herodotus, using *Hestia* data, see: E. T. E. Barker, S. Bouzarovski, C. B. R. Pelling, and L. Isaksen, ‘Extracting, investigating and representing

physical features envelop anthropogenic constructs like territories, which act as anchors for the geography of linkages between social formations constructed by Herodotus.

A whole range of other basic network maps can be quickly and easily generated that vary in nature according to the place type one chooses to focus on (settlement, territory, or natural feature), the general scene in a particular book, the connections enjoyed by any one specific place, those places that enjoy the most number of connections (the central ‘hubs’ in a network), or those networks that have the strongest connections (the thickest ‘edges’). Thus the database can be used not only to gain a general sense of a broad network culture represented in Herodotus, but also to generate a series of different types of network representations depending on the queries that one asks of the data. It is worth reiterating, however, that none of these maps are sufficient in and of themselves: rather, they are there to prompt new questions and provoke further investigation. In this way, all these GIS-generated maps, both the simple ‘places in isolation’ and the more complex networks, should be regarded as *complementing and preceding* rather than replacing close textual analysis.

5. Bringing Herodotus’ world into everybody’s home

The foregoing sections have discussed the possibilities for GIS spatial and topological analysis of places within Herodotus’ *Histories*. At the same time, our experiences have found that using GIS has demanded a level of expertise higher than that enjoyed by most of the team – a point not often acknowledged in the literature on humanities GIS. Limitations in its usability, however, were for us an important issue, since we have been determined from the outset to make our data, methodologies, technologies, and outcomes open for general use and reuse.³² GIS, in our minds, did not have the widespread applicability that we were looking for.

In order, then, to start experimenting with public dissemination, we decided to expose the PostGIS data as Keyhole Markup Language (KML), a markup format that can be read by a variety of webmapping applications including the hugely popular Google Earth. This was achieved by installing the Open Source GeoServer, which serves spatial data in a variety of web-friendly formats simultaneously, including KML, SVG, WMS, WFS, and PDF. Significantly, it is automatically readable over the web on any machine that has Google Earth installed, and, since the link is a network link rather than a static KML file, any changes that we made to the *Hestia* database would result in automatic (and more-or-less instantaneous) updating of the data in the viewing application without any need on the user’s part to do anything.

geographical concepts in Herodotus: the case of the Black Sea’, in *The Bosphorus: Gateway between the Ancient West and East*, eds. G. R. Tsetskhladze et al. (Oxford 2013) 7-17.

³² In part this was in response to the concern expressed by our original sponsor, the AHRC, that their funded projects be disseminated as widely as possible. But we also hold that practising openness has proven overwhelmingly beneficial for our research by encouraging potential collaborators to get in touch. We note too that our own open approach has been largely possible because of *Perseus*’ relatively liberal Creative Commons licence, which directly encourages further use and wider dissemination. Within the digital world, overly restrictive licencing (such as No-Derivatives clauses) can be almost as pernicious to scholarship as the failure to make the data available in the first place.

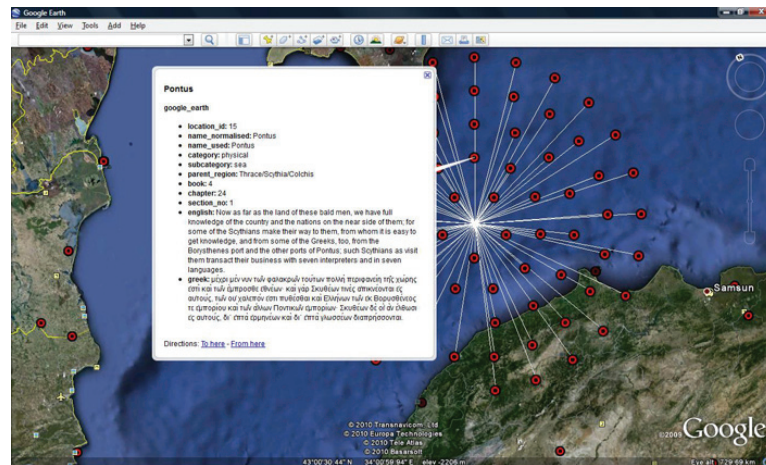


Figure 11: The *Histories* in Google Earth, highlighting one of the passages (each indicated by a red circle) in which Herodotus talks about the Pontus

http://hestia-geo.open.ac.uk:8080/geoserver/wms/kml?layers=hestia:google_earth

With this ‘Herodotus geodata’, users are able to construct ‘mashups’ of visual and textual data.³³ For example, since all places are linked to entries in the database, by simply clicking on a particular location in Google Earth, users are able to bring up a balloon box containing Herodotus’ text (in both English and Greek) for that particular location for every occasion when it is mentioned in the narrative (Figure 11). It is this part of the project that potentially has the greatest impact, with its capacity to reach beyond an academic environment and bring the world of an ancient historian into peoples’ homes. In the case of Herodotus’ *Histories* this might involve linking different places that are mentioned with each other in the same stretch of narrative in chronological order, so that it would be possible to follow the ‘journey’, say, of historical agents within the text. One could, for example, imagine following Xerxes’ passage into Greece from the ‘down-on-the-earth’ perspective of one of his myriad of troops.

While Google Earth is an excellent vehicle for enabling users of all kinds to explore the geography of Herodotus’ *Histories*, like GIS it is less effective at displaying the changing points of geographical reference over the course of the narrative. The inherent danger in this kind of technology is that one forgets that there is a text to be read or, in Herodotus’ day, to be heard – and reading or listening is an experience that happens in time. Therefore in collaboration with Nick Rabinowitz, developer of Timemap.js,³⁴ we

³³ As the account in Wikipedia puts it, ‘a mashup in web development is a web page, or web application, that uses content from more than one source to create a single new service displayed in a single graphical interface’. For more information, see: [http://en.wikipedia.org/wiki/Mashup_\(web_application_hybrid\)](http://en.wikipedia.org/wiki/Mashup_(web_application_hybrid)) [last accessed on 15/11/2013].

³⁴ For Timemap.js, see n.10 above. The Herodotus TimeMap has since been superseded by GapVis. See section 6 below.

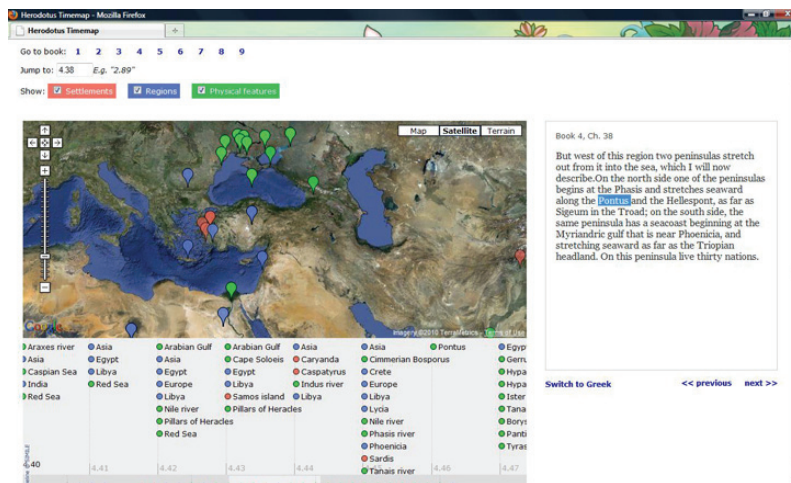


Figure 12: Herodotus TimeMap. Reading text, map, and (narrative) time alongside each other. <http://www.nickrabinowitz.com/projects/timemap/herodotus/basic.html>

have created a web-based interface combining a map, a ‘narrative timeline’, and a reading pane, allowing the reader to view locations as they are mentioned in the text. For example, Figure 12 shows all the places mentioned by Herodotus in or near 4.38, divided according to settlement, territory or physical feature. The central Google Map displays geographical concepts roughly collocated in the narrative; the timeline below depicts ‘narrative time’ represented as the sequence of book sections, each listing geographical references cited within them; to the right, a reading window presents the text of the current chapter, which can be toggled between English and Greek. Navigational elements allow the user to page through the text or jump to a specific section, while clicking on a location in the map or timeline will move the text to the relevant chapter, highlighting the selected *toponym*. Markers on the map fade to transparency as the narrative progresses beyond their mention, mimicking their gradual disappearance from the scope of the reader’s attention.

This ‘timemap’ of Herodotus’ world represents a novel use of the Timemap.js library, which was designed to combine maps with historical, rather than narrative, timelines. The JavaScript library, which integrates the Google Maps API with the SIMILE Timeline API,³⁵ facilitates the loading of data simultaneously onto both a map and timeline, with a separate data structure to manage cross-references and interactions between the two visual representations of a given record. The result is an online interface in which user interaction with either the map or the timeline can be reflected in both elements; for example, clicking an item on the timeline may bring up an informational window on the map, and markers on the map may be hidden when their corresponding timeline item is out of view. To customize Timemap.js to the Herodotus Timemap application, a custom set-up script was required to convert the date-based chronological timeline into one displaying the narrative sequence of book chapters and subsections. To avoid the need for

³⁵ SIMILE project Timeline web widget: <<http://www.simile-widgets.org/timeline>>.

revisions to the core architecture of the SIMILE Timeline, this was accomplished through the relatively simple expedient of mapping chapters to an arbitrary sequence of dates, then using that mapping to determine the timeline labels and item positions.³⁶ The other technical challenge was to find a method for loading a large number of data items without overloading the browser-based interface. Under normal circumstances, a Google Map cannot display more than a few hundred items without significant performance issues – and our dataset included over 4,000 place references. Our solution, which addressed both performance and load-time concerns, was to segment both chapter text and location data into static JSON³⁷ files of manageable size, which could then be loaded ‘on demand’ as the user progressed through the narrative.

The Timemap is less an analytical tool than an enhanced interface for engaging with Herodotus’ text as a reader. Unlike several of the data visualization approaches discussed above, it displays comparatively little information at any given time, focusing attention instead on the reading experience of the *Histories* and the relationship of geographical concepts to their place in the narrative. Our hope is that the rich visual detail that the Timemap adds to the text can offer value at the level of both the novice reader, who may benefit from a clear depiction of the locations under discussion, and the researcher, for whom the interface may help to identify quickly places co-occurring in the narrative and highlight chapters where Herodotus’ use of geographical reference is particularly focused or wide-ranging. In fact, by trying to map as closely as possible the reading experience, we hope that the Herodotus Timemap may have the additional utility of facilitating research into the ways in which geographical concepts in the *Histories* undergo change over time.

All-in-all, these different technologies have allowed us to develop unprecedented means of disseminating the world view of an ancient historian as well as providing a glimpse of the potential for digital resources to help expose the spatial data for examination and ask new questions of it.

6. Towards a digital analysis of ancient places?

Applying new technologies to a dataset such as that of an ancient Greek historian inevitably raises (at least) as many questions as it resolves. Issues that have arisen during this project relate to data collection, management and aggregation, accuracy and precision, expertise required to use the tools, above all, the ‘so what?’ question. Indeed, our experience of using digital resources and developing an automated process for the querying and visualization of spatial data in Herodotus has demonstrated beyond doubt the need to adopt plural approaches that make use of, but are not dependent on, the technology. In short, we conclude that the employment of the digital medium can greatly enhance our understanding of spatial concepts in Herodotus *provided that* it is accompanied by a close engagement with the text at all stages of the analysis. Ultimately, it is this need to marry ‘quantitative’ and ‘qualitative’ approaches that we regard as one of the most interesting implications of the initial steps that we have taken.

³⁶ Besides, many of the dates associated with events or characters within the *Histories* are known only within very rough limits, particularly those that occur before the fifth century BC.

³⁷ JSON (JavaScript Object Notation) is a ‘lightweight data-exchange format’:
<<http://www.json.org/>>.

Of course *Hestia* is not alone in pioneering approaches to adopting digital technologies in classical studies. A similar project in the field of archaeology is *Open Context*, an Open Repository of archaeological excavation data developed and maintained by the Alexandria Archive and UC Berkeley.³⁸ In 2010, a funding call from Google, under its Digital Humanities Award program, sought to enable researchers to ‘apply quantitative research techniques for answering questions that require examining thousands or millions of books’. This presented an opportunity to researchers on both *Open Context* and *Hestia* to pursue the common goal of tying classical resources together through referencing places.³⁹ The central principle behind the GAP (Google Ancient Places) project, identified while developing the *Hestia* Timemap, is that places referenced in narrative texts generally cluster together to maintain narrative coherency. In other words, given a set of *toponyms* with multiple possible identifications, the set of identifications with the shortest overall path between them is likely to be correct. We can further weight the influence of each *toponym* on our decision by the number of possible locations to which it could refer. Somewhat counter-intuitively, this means that small, obscure places with unusual names are much better guides to location than well-known places with many namesakes.⁴⁰

Since 2011, the GAP team have been working in conjunction with the *Pleiades* project⁴¹ to associate the local identifiers used by the *Hestia* project with *Pleiades* Uniform Resource Identifiers (URIs) for each place to which multiple names (*toponyms*), locations (such as spatial co-ordinates), and categories (like ‘settlement’) can be assigned. Not only does this make it much easier to handle the problem of synonymy, but it also means that, once an identification is made, it can be permanently fixed with a non-ambiguous identifier.⁴² At a stroke this process renders digital data much more valuable, for now various kinds of enquiry can be asked of the data and various kinds of information can be brought together: it will be possible to find out about ancient places not only in texts but in other resources, including tables, databases, maps and images.⁴³ In addition, GAP is experimenting with ways

³⁸ For more on *Open Context* see: <<http://opencontext.org/>>.

³⁹ Google Research Blog: ‘Our commitment to the digital humanities’: <http://googleresearch.blogspot.co.uk/2010/07/our-commitment-to-digital-humanities.html> [last accessed on 15/11/2013].

⁴⁰ For more information about the GAP project, including regular posts describing the work that has been carried out, see: <<http://googleancientplaces.wordpress.com/>>. The GAP team have also produced a new reading interface called GapVis (<http://gap.alexandriaarchive.org/gapvis/index.html>). This platform provides an initial landing page showing all the places in a text visualised on a map and as a histogram; a ‘reading page’ for viewing the places alongside the narrative; and a ‘place page’ for focusing on information about a particular place, such as those places mentioned most frequently with it (see figure 13).

⁴¹ The *Pleiades* project (<<http://pleiades.stoa.org/>>) is in the process of digitizing the *Barrington atlas of the classical world*. The fact that *Pleiades*’ coverage for Herodotus’ place names had not been available at the time of the initial *Hestia* project demonstrates the rapidly changing digital environment.

⁴² Or URIs (Uniform Resource Identifiers): <http://en.wikipedia.org/wiki/Uniform_Resource_Identifier>.

⁴³ Led by Elton Barker and Leif Isaksen of *Hestia*, Pelagios aims to make sense of the sea of data related to ancient world places. Funded by JISC in two consecutive programmes (Geospatial

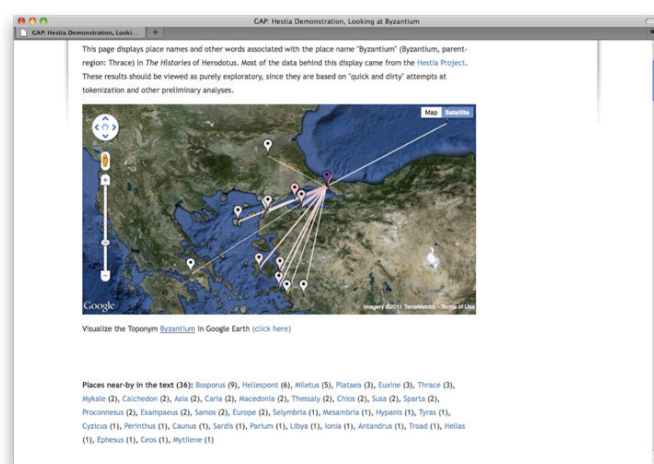


Figure 13: 'Place' page view of the GapVis platform, showing the places most frequently mentioned with 'Byzantium' in the *Histories*

of visualizing and analyzing spatial data that does not demand specific expertise in using specialist applications, such as GIS. Figure 13, for example, depicts a network analysis map of Byzantium's network in Herodotus (based on the co-reference of places in the text) in KML, which is the kind of map that any users will be able to generate and explore for themselves.

In conclusion, *Hestia* is not an end product, but just the beginnings of what promises to be a radical reconfiguration of the research (and teaching) process. We believe that the technologies described here, when used together and in conjunction with traditional scholarship, will help not only facilitate innovative lines of enquiry and lead to new insights, but also bring Classics to a whole a new audience. Digital Classics is already starting to move beyond the era of stand-alone projects, and the tools and data described here are already openly available. It is now time for the community to build upon this material, opening it up both to classicists and to the world at large.

Elton Barker (*Open University*) Elton.Barker@open.ac.uk
 Leif Isaksen (*University of Southampton*) L.Isaksen@soton.ac.uk
 Nick Rabinowitz nick.rabinowitz@gmail.com
 Stefan Bouzarovski (*University of Manchester & University of Gdansk*)
 stefan.bouzarovski@manchester.ac.uk
 Chris Pelling (*Christ Church, Oxford*) chris.pelling@chch.ox.ac.uk

Engagement and Community Outreach programme; Resource and Discovery programme) Pelagios (meaning 'of the sea') is establishing generic processes and practices for referencing ancient places in online documents. For more information, please see the project blog: <<http://pelagios-project.blogspot.com/>>. As this publication goes to press, Pelagios is entering its third phase of development. Now with the support of the Andrew W. Mellon Foundation, and led by Isaksen, Pelagios will annotate, link, and index place references in digitized Early Geospatial Documents that use written or visual representation to describe geographic space prior to 1492.